Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-429-SC1, 2018

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Quantifying new water fractions and transit time distributions using ensemble hydrograph separation: theory and benchmark tests" *by* J. Kirchner

N. Rodriguez

nicolas.rodriguez@list.lu

Received and published: 18 October 2018

General comments:

This paper has been really pleasant to read. The quality of the writing, of the figures, and of the mathematics is really high. The amount of meaningful explanations is impressive. I think the paper will have a strong impact on the isotope hydrology community. I think we need more approaches such as this one in tracer hydrology. That being said, I would like to mention two things that could be discussed.

The first one is more context about travel time modeling. While the introduction and

Printer-friendly version

Discussion paper

the discussion compare well this ensemble approach to the traditional hydrograph separation, a large part of the paper also deals with determining travel time distributions (TTDs). Yet, only little is said about travel time modeling, especially in the introduction. I think it should be mentioned that the ensemble approach deals with a current need in isotope hydrology to have more data-driven approaches and non-parametric TTDs. I think this is exactly what the proposed approach brings compared to already existing approaches, but not more. Unlike reviewer n°1, I believe we should not try to formally compare methods which have different purposes. This approach calculates only the streamflow average TTD suggested by the tracer data, without assuming its shape. This is novel and important. Yet the proposed method can only be used for the period covered by training data (i.e. "backwards"), and for streamflow only. Using StorAge Selection (SAS) functions with assumed shapes allows one to obtain the time-varying TTDs at every moment and in every flux (backwards), and the Residence Time Distributions (RTDs). But more importantly, SAS function allow one to simulate other time-varying solute fluxes (e.g. Benettin et al., 2015) with the calibrated model in a forward way (even outside the period covered by training data). Note that a model based on SAS functions can consist of just a handful of parameters (e.g. Benettin et al. 2017) which makes it really competitive. Yet, I also agree that there are clear limitations in approaches based on SAS functions.

My second comment relates to the potential limitations of the proposed approach. I think that all the choices made to derive the mathematical solutions were presented as if they are the best choices for any tracer data set, or the only choices possible. This may not be true in all cases. The discussion would benefit from an objective assessment of the problems that could occur when trying to apply the approach to real tracer data. In my opinion, this ensemble approach will be accurate only for the left tail of the TTDs, while it truncates (cf. equation 30) older ages. This is a critical problem in travel time modeling in general (Stewart et al., 2012; Stewart & Morgenstern, 2016). It is already mentioned in the reply to reviewer n°1, but I think it should be clearly written in the discussion as well.

Specific comments:

(1) To give more context in the introduction you could mention and describe briefly the common methods to estimate TTDs, namely the Lumped Parameter Models (e.g McGuire & McDonnell, 2006, and references therein), flux tracking in conceptual models (e.g. Hrachowitz et al., 2013), SAS functions applied to a single control volume (e.g. Benettin et al., 2017), and particle tracking in distributed models (e.g. Davies et al., 2013; Danesh-Yazdi et al., 2018). Doesn't the ensemble approach answer the need to have alternatives to these methods, which all need to assume an underlying model for water transport?

(2) P7, L11-13: Least squares regression means that any real data set with "outliers" (which may just be tracer values one did not expect) is likely to adversely affect the results from the ensemble approach, as it is suggested here. Same for the least squares solution in equation 38. This is in my opinion one of the limitations of the proposed method. This should perhaps be mentioned in the discussion. Can iteratively reweighted least squares or another robust regression technique be used instead? I agree that this approach assumes no model for the transport of tracers, yet it does assume a model for the errors between the regression and the measurements (i.e. the residuals). This is similar to the choice of an objective function in traditional model calibration, and deserves attention. For example, commonly used assumptions about streamflow residuals were shown to be often violated, because of autocorrelation, non-normality, and heteroscedasticity (Schoups & Vrugt, 2010). Are the tracer residuals in this work likely to show non-normality, autocorrelation, and heteroscedasticity as well? Although CQ(j) and Cnew(j) are both "normalized" by subtracting CQ(j-1), there could be autocorrelation of higher order than just 1. How does the variance of errors change with larger flashy events?

(3) P12, L19-20 All the benchmarking is done for a catchment without evapotranspiration. This points to a more general concern with the ensemble approach. No assumption is made explicitly about what happens to the tracer masses between precipitation

and streamflow. This means that the method may try to find direct "connections" (in a loose statistical sense here) between tracer inputs from the past and current tracer fluctuations in the stream. Intermediary (unconsidered) processes may still be important to explain the transformation from one to the other. I especially think of processes affecting the lumped catchment tracer mass balance, which is an expression that was not considered in the approach. In that regard, how are the results expected to change if ET is actually used in the benchmark tests? Is the approach robust for real catchments where ET can be a major part of the water balance? Here I am not considering the effects of fractionation which were already dealt with, but the selective removal of certain tracer masses (associated with particular ages, i.e. different soil/groundwater mixtures) by ET, which will hence not be available for streamflow.

(4) P20, L4-6: These estimates seem to differ as much as 50% from the known values for the damped catchment and weekly data on figure 4. Many tracer data sets are at weekly resolution and come from "damped" catchments (e.g. Tetzlaff et al., 2009; Pfister et al., 2017). Data-driven approaches are by nature highly sensitive to the quality of data (e.g. variability, resolution, and measurement uncertainty). The proposed approach could thus show some limitations due to its strict data needs in some cases. This could be mentioned in the discussion.

(5) P20, L19-20: A weekly sampling routine is likely to contain more "baseflow" samples which reflect older water contributions. This results in an underestimation of QFnew as shown here. Yet how can the fraction of new water with respect to discharge be underestimated while the fraction of new water with respect to precipitation is overestimated? These quantities refer to the same mass of water in streamflow.

(6) P28, L2: Here it is assumed that values of Cp and CQ at all times corresponding to indexes j or j-k are known (except a few, which require the solutions proposed in 4.2 and 4.3). In practice it is very likely that the sampling interval is irregular, such that there is not a perfect correspondence between measurement times, and required times indexed by j or j-k. Any recommendation on how to best adjust the measurement

Printer-friendly version

Discussion paper

time series so that these terms are defined properly would be welcome. Similarly, the method requires the same number of measurements in precipitation and in streamflow. How could we deal with this in various research catchments as this is often not the case?

(7) P29, L12: Here it is assumed that the most recent precipitation events have more weight in the current tracer fluctuations in the stream than older inputs. This is implicitly reflected by the truncation of the sum in equation 30. This is also reflected by the estimated travel times that mostly stay below a few months. Although this assumption about tracer contributions is likely to be valid, catchment travel times are known to be generally in the order of magnitude of a few years and even decades (McGuire & McDonnell, 2006). This is all the more true when age estimates are based on tritium measurements (Stewart et al., 2010). Would the ensemble approach be robust in catchments where streamflow is volumetrically dominated by water older than a few months?

(8) P29, L12: Regarding the linear algebra, how large can the truncation index m be in practice, given that computationally intensive large matrix operations are carried out? This is especially of interest since the matrices grow with the number of measurements in both dimensions, while m needs to be as large as possible for the ensemble method to work well. In my opinion, the discussion should encourage the reader to consider if this approach shows limitations for his/her considered travel times, which may be up to a decade. Can this approach go beyond the left-hand tail of the TTDs or is it limited to the left tail?

(9) P33, L8: Over several years of data, doesn't neglecting 1 mm of precipitation per day sum up to a large value? It could be useful to include some discussion on the effect of that threshold on the results. Are the results highly sensitive to that choice or not?

(10) P39, Figure 11: Deviations between the benchmark TTDs and the estimated ones

are visible here. How could these deviations be described more quantitatively to be more objective?

(11) P40, Figure 12: It looks like the uncertainties are larger for the TTDs which shape is not a classical “L” anymore. The explanation given here is that the effective sample size neff (equation 13) is small because of tracer autocorrelation. Can we not say that an autocorrelation in the tracer time series is universal, as well as a shape of the TTD far from a simple “L”? This seems related to the issue described in comment (2).

(12) There are not many data-driven methods that can yield non-parametric TTDs, which explains why this new approach is really beneficial to estimate TTDs. Yet, I believe that the problem solved here is somewhat similar to what Turner et al. (1987) solved as well, using Kalman filtering approaches (see the parallel between equations 30 and 31 here, and their equations 1 and 2) (see also Turner and McPherson, 1990). They unfortunately did not detail the math behind their approach. Nevertheless, their work present time-varying average transit times, including uncertainties, also derived without assumptions on the shape of the TTDs. This is worth mentioning and comparing to the presented approach in the discussion. Furthermore, Klaus et al., (2015) also presented a data-driven approach that could be worth mentioning and briefly comparing to the presented one. Finally Kim et al. (2016) could estimate not only TTDs but also SAS functions from artificial tracer data (of course under well controlled lab conditions). Their work is worth mentioning because they are able to distinguish the “external” variability of travel times from the “internal” one, unlike the ensemble approach presented here.

(13) P54, L30-32: Doesn’t this mean that m should be set bigger?

Small technical comments:

(14) Figure 1 & 4: the colors are not consistent between the legend and the lines in the lower subplots.

References: Benettin, P., S. W. Bailey, J. L. Campbell, M. B. Green, A. Rinaldo, G. E. Likens, K. J. McGuire, and G. Botter (2015), Linking water age and solute dynamics in streamflow at the Hubbard Brook Experimental Forest, NH, USA, Water Resour. Res., 51, 9256–9272, doi:10.1002/2015WR017552.

Benettin, P., C. Soulsby, C. Birkel, D. Tetzlaff, G. Botter, and A. Rinaldo (2017), Using SAS functions and high-resolution isotope data to unravel travel time distributions in headwater catchments, Water Resour. Res., 53, 1864–1878, doi:10.1002/2016WR020117.

Danesh-Yazdi, M., Klaus, J., Condon, L. E., & Maxwell, R. M. (2018). Bridging the gap between numerical solutions of travel time distributions and analytical storage selection functions. Hydrological Processes, 32, 1063–1076. https://doi.org/10.1002/hyp.11481

Davies, J., K. Beven, A. Rodhe, L. Nyberg, and K. Bishop (2013), Integrated modeling of flow and residence times at the catchment scale with multiple interacting pathways, Water Resour. Res., 49, 4738–4750, doi:10.1002/wrcr.20377.

Hrachowitz, M., Savenije, H., Bogaard, T. A., Tetzlaff, D., & Soulsby, C. (2013). What can flux tracking teach us about water age distribution patterns and their temporal dynamics? Hydrology and Earth System Sciences, 17(2), 533–564. https://doi.org/10.5194/hess-17-533-2013

Kim, M., Pangle, L. A., Cardoso, C., Lora, M., Volkmann, T. H. M., Wang, Y., et al. (2016). Transit time distributions and StorAge Selection functions in a sloping soil lysimeter with time-varying flow paths: Direct observation of internal and external transport variability. Water Resources Research, 52, 7105–7129. https://doi.org/10.1002/2016WR018620

Klaus, J., Chun, K. P., McGuire, K. J., & McDonnell, J. J. (2015). Temporal dynamics of catchment transit times from stable isotope data. Water Resources Research, 51, 4208–4223. https://doi.org/10.1002/2014WR016247

McGuire, K. J., & McDonnell, J. J. (2006). A review and evaluation of catchment transit time modeling. Journal of Hydrology, 330(3–4), 543–563. https://doi.org/10.1016/j.jhydrol.2006.04.020

Pfister L, MartínezâĂŘCarreras N, Hissler C. et al. Bedrock geology controls on catchment storage, mixing, and release: A comparative analysis of 16 nested catchments. Hydrological Processes. 2017. https://doi.org/10.1002/hyp.11134

Schoups, G., and J. A. Vrugt (2010). A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and nonâĂŘGaussian errors, Water Resour. Res., 46, W10531, doi:10.1029/2009WR008933.

Stewart, M. K., & Morgenstern, U. (2016). Importance of tritium-based transit times in hydrological systems, WIREs Water 2016, 3:145–154. doi:10.1002/wat2.1134

Stewart, M. K., Morgenstern, U., & McDonnell, J. J. (2010). Truncation of stream residence time: how the use of stable isotopes has skewed our concept of streamwater age and origin, Hydrological Processes 24, 1646-1659, doi:10.1002/hyp.7576

Stewart, M. K., Morgenstern, U., McDonnell, J. J., & Pfister, L. (2012). The 'hidden streamflow' challenge in catchment hydrology: A call to action for stream water transit time analysis. Hydrological Processes, 26(13), 2061–2066. https://doi.org/10.1002/hyp.9262

Tetzlaff, D., Seibert, J., McGuire, K. J., Laudon, H., Burns, D. A., Dunn, S. M., et al. (2009). How does landscape structure influence catchment transit time across different geomorphic provinces? Hydrological Processes, 23(6), 945–953. https://doi.org/10.1002/hyp.7240

Turner, J. V., & Macpherson, D. K. (1990). Mechanisms Affecting Streamflow and Stream Water Quality: An Approach via Stable Isotope, Hydrogeochemical, and Time Series Analysis. Water Resources Research, 26(12), 3005-3019.

Turner, J.V., Macpherson, D.K., Stokes, R.A., 1987. The mechanisms of catchment

flow processes using natural variations in deuterium and oxygen-18. Journal of Hydrology 94, 143–162.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-429, 2018.

Printer-friendly version

Discussion paper